KONINKRIJK DER NEDERLANDEN

Bureau voor de Industriële Eigendom

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Hierbij wordt verklaard, dat in Nederland op 26 juli 2002 onder nummer 1021160,

ten name van:

MULTIGEN HOLDING S.A.

te Luxembourg, Luxemburg

een aanvrage om octrooi werd ingediend voor:

"Method of determining the copy number of a nucleotide sequence",

en dat de hieraan gehechte stukken overeenstemmen met de oorspronkelijk ingediende stukken.

Rijswijk, 5 augustus 2003

De Directeur van het Bureau voor de Industriële Eigendom,
voor deze,

Mw. M.M. Enhus

ABSTRACT

The invention relates to a method of determining of accurately determining the copy number of a nucleotide sequence I in a sample using an amplification technique, such as PCR. According to the present invention, a second nucleotide sequence II is also measured and calibration curves for each are made, from which the relative copy number CN can be determined.

NL 45369-A1/ho

Method of determining the copy number of a nucleotide sequence

The present invention relates to a method of determining the copy number of a nucleotide sequence I in a sample using an amplification technique, said method comprising the steps of

1) adding nucleotides, primers, polymerase and any further reagents, if any, required for the amplification technique used to the sample,
2) performing one or more amplification cycles to amplify the nucleotide sequence for which the copy number has to be determined.

Most eukaryotic diploid cells contain two copies of a single gene; one on each chromosome of a pair of chromosomes. The chromosomes of a pair of chromosomes being derived from each parent, the genes may be different and, for example, one of them may result in a abnormal protein. Thus, the number of functional genes is not necessarily 2 in an eukaryote. While often genes are present in one copy per chromosome of a particular pair of chromosomes, some genes are present in multiple copies, for example in tandem repeat sequences. Another exception to the general rule of 2 per cel is mitochondrial DNA. A cell contains many mitochondria, the number being dependant on the type of cell. But even for a particular cell type, the number of mitochondria may vary. Typical numbers are between 100 and 1000 mitochondria per cell, and each mitochondrion contains several copies of mitochondrial DNA. The copy number is not necessarily equal to larger than 2 per cell. Some nucleotide sequences are very rare among cells (despite being of one and the same subject, such as a human being). This is, for example, after gene rearrangement. This is, for example, the case with antibody producing cells (B-lymphocytes) or receptor-carrying T-lymphocytes. Of a large number of lymphocytes, only a few will contain a particular nucleotide sequence defining the variable region of a particular antibody (or of the T-cell receptor), capable of recognizing a particular antigen. In the art, a need exists

to reliably determine the copy number of a nucleotide sequence, which may comprise the nucleotide sequence of a gene or part thereof. A method according to the preamble is known in the art. In particular Douek et al (Nature 396, pp. 690-695 (1998)) describe a method for detecting the products of the rearrangements of T-cell receptors (TREC) using a semi-quantitative assay. For determining the amount of TREC in a given sample, a known amount of a DNA competitor are prepared. Then, an amount of sample DNA containing the nucleotide sequence to be determined are added to the tube. A PCR amplification reaction is carried out in the presence of radiolabeled deoxynucleotide. Subsequently, the resulting amplification products are run on a gel to separate the sample DNA PCR product from the competitor DNA product. After autoradiography, the amount of nucleotide sequence to be determined is calculated using densitmetric analysis from the ratio between a band of competitor and a band of the sample DNA. The result is expressed as the number of copies of TREC per microgram total DNA. To achieve an acceptable accuracy, 4 tubes containing scalar amounts of competitor DNA are used, to which fixed amounts of sample DNA are added.

The disadvantage of this method is that when DNA is extracted from cells, it must be assumed that this is all the DNA present in the cells. That is, it is assumed that no cell escaped lysis and all DNA present in the cells was extracted and isolated. This is not necessarily the case. Another disadvantage of this method is that it is sensitive to differences in amplification efficiency.

The object of the present invention is to provide a method capable of reliably determining the copy number of a nucleotide sequence even if it is present in extreme amounts, such as lots of copies per cell or only few copies per many cells. In addition, an object of the present invention is to provide a method which has reduced sensitivity to the efficiency with which DNA has extracted from the cells containing a nucleotide sequence I for which the copy number has to be determined.

To this end, the method according to the present in-

vention is characterized in that the sample contains a chromosomal second nucleotide sequence II, and

- the first nucleotide sequences I is amplified,
- the second nucleotide sequences II is amplified,
- a third nucleotide sequence I′ corresponding to the first nucleotide sequence I is amplified at various dilutions, and
- a fourth nucleotide sequence II′ corresponding to the second nucleotide sequence II is amplified at various dilutions,

where the ratio of the concentrations of nucleotide sequence I′ and II′ is known; the amplifications of the third and fourth nucleotide sequences I′ and II′ at various dilutions allows standard curves SC_i with i being I or II to be made, the concentrations of I and II are determined by using the respective standard curve SC_i , and the relative concentrations allows the relative copy number CN of sequence I (versus nucleotide sequence II) to be determined using the formula

$$CN = \frac{[I]_{SCI'}}{[II]_{SCII'}}$$

where

CN is the relative copy number of I over II in the sample;

$[I]_{SCI'}$ is the concentration of I determined using standard curve SCI′; and

$[II]_{SCII'}$ is the concentration of II determined using standard curve SCII′.

The nucleotide sequence concentrations can be determined using any suitable methods known in the art. Such a method may comprise radioactive methods, the use of electrophoretic separations and quantification using, for example, densitometry etcetera. While specific reference is made to standard curves, it goes without saying that this can be done using computational methods without an actual graph being made. This is fully within the scope of the present inven-

tion. Generally, all amplifications will be performed separately (i.e. in an individual amplification container such as a well of a microtiter plate), and preferrably substantially at the same time. The method according to the invention is not only highly accurate, but it is also very efficient if performed for mulitple samples.

The number of amplification cycles are not necessarily the same for I and II, but they are the same for a) I and I'; and for II and II'.

According to a preferred embodiment the absolute copy number is determined by multiplying the copy number CN by the absolute copy number of sequence II per cell.

For several nucleotide sequences II the number of copies of per cell is known. An example is, for example, the gene coding for heat shock protein 70, or Fas Ligand (CD178), which are known to be present with two copies per cell (i.e. the absolute copy number of hsp 70 = 2). Many nucleotide sequences of genes are very suitable because they generally are present in a known number of copies in every cell of the species from which the DNA is derived. The efficiency with which DNA material is extracted from the cells is not important (although, in case nucleotide sequence I is on a different molecule as nucleotide sequence II, it is important that they are extracted with the same efficiency). Hence, this embodiment allows determination of the absolute copynumber of the nucleotide sequence I per cell.

According to a preferred embodiment, at least one of the third nucleotide sequence I' and fourth nucleotide sequence II' resides on a vector.

When present on a vector, it is easy to obtain that particular nucleotide sequence in desired quantities. It is also very easy to determine the DNA concentration and hence the copy number of the nucleotide sequence per volume. The vector may be any vector known in the art, such as a plasmid, a cosmid, a virus etc. If the third nucleotide sequence I' resides on first vector and the fourth nucleotide sequence II' resides on a second vector, the vectors can be used (or mixed) at any desired ratio to accommodate expected differ-

ences in copy number in the sample.

A vector carrying the third nucleotide sequence I' preferably contains at least two different third nucleotide sequences I' for measuring two different first nucleotides I.

In other words, a single vector, requiring its concentration to be determined only once, can carry multiple third nucleotide sequences I', which allows, for example, the copy numbers of many different genes to be determined.

It is highly prefered that the third nucleotide sequence I' and fourth nucleotide sequence II' reside on the same vector.

Thus the ratio is constant and exactly known (for example 1:1). This allows for the most accurate measurements possible.

Preferrably, the first nucleotide sequence I is the same as the third nucleotide sequence I'.

This strongly reduces errors due to differences in amplification efficiencies between I and I'. Nevertheless, small differences in nucleotide sequence are generally allowed, although changes at locations where the probe used for detecting the concentration of the nucleotide sequence are best avoided.

Similarly, it is preferred that the second nucleotide sequence II is the same as the fourth nucleotide sequence II'.

According to a preferred embodiment, the vector contains at least two different nucleotide sequences I'.

According to a highly preferred embodiment, the amplification reaction is followed, preferably spectrophotometrically.

Such a spectrophotometrical method is known in the art. More specifically, internal probes can be used which allow for real time measurements, for example real time PCR. Internal probes are known in the art, and are disclosed by, for example, Winer et al (Anal. Biochem 270, pp. 41-49 (1999)). Measurements can be done either continuously, or after finishing an amplification cycle.

While the present invention is described with refer-

ence to DNA, the present invention also applies to the determination of the number of RNA sequences present in a cell. Use can be made of methods known in the art to multiply RNA, for example by preparing cDNA. This application does not attempt to teach in interested layman how to become a person skilled in the art, for which reason the layman is referred to general text books and in particular to a proper university to learn the required techniques that a person skilled in the art knows how to apply these techniques to work the present invention.

CLAIMS

1. Method of determining the copy number of a nucleotide sequence I in a sample using an amplification technique, said method comprising the steps of

1) adding nucleotides, primers, polymerase and any further reagents, if any, required for the amplification technique used to the sample,

2) performing one or more amplification cycles to amplify the nucleotide sequence I for which the copy number has to be determined;

characterized in that the sample contains a chromosomal second nucleotide sequence II, and

- the first nucleotide sequences I is amplified,
- the second nucleotide sequences II is amplified,
- a third nucleotide sequence I′ corresponding to the first nucleotide sequence I is amplified at various dilutions, and
- a fourth nucleotide sequence II′ corresponding to the second nucleotide sequence II is amplified at various dilutions,

where the ratio of the concentrations of nucleotide sequence I′ and II′ is known; the amplifications of the third and fourth nucleotide sequences I′ and II′ at various dilutions allows standard curves SC_i with i being I or II to be made, the concentrations of I and II are determined by using the respective standard curve SC_i , and the relative concentrations allows the relative copy number CN of sequence I (versus nucleotide sequence II) to be determined using the formula

$$CN = \frac{[I]_{SCI'}}{[II]_{SCII'}}$$

where

CN is the relative copy number of I over II in the sample;

$[I]_{SCI'}$ is the concentration of I determined using standard curve SCI'; and
$[II]_{SCII'}$ is the concentration of II determined using standard curve SCII'.

2. Method according to claim 1, **characterized** in that the absolute copy number is determined by multiplying the copy number CN by the absolute copy number of sequence II per cell.

3. Method according to claim 1 or 2, **characterized** in that at least one of the third nucleotide sequence I' and fourth nucleotide sequence II' resides on a vector.

4. Method according to claim 3, **characterized** in that the vector contains at least two different third nucleotide sequences I' for measuring two different first nucleotides I.

5. Method according to claim 4 or 5, **characterized** in that the third nucleotide sequence I' and fourth nucleotide sequence II' reside on the same vector.

6. Method according to any of the preceding claims, **characterized** in that the first nucleotide sequence I is the same as the third nucleotide sequence I'.

7. Method according to any of the preceding claims, **characterized** in that the second nucleotide sequence II is the same as the fourth nucleotide sequence II'.

8. Method according to any of the preceding claims, **characterized** in that the amplification reaction is followed, preferably spectrophotometrically.